BAB II

TINJAUAN PUSTAKA

2.1 Peramalan

Peramalan adalah proses perkiraan besarnya atau jumlah sesuatu pada waktu yang akan datang berdasarkan data pada masa lampau yang dianalisis secara ilmiah khususnya menggunakan metode statistika [8]. Hal ini dapat dilakukan dengan melibatkan pengambilan data di masa lalu dan menempatkannya ke masa yang akan datang dengan suatu model matematis. Teknik peramalan terbagi menjadi dua kelompok yaitu teknik kualitatif dan teknik kuantitatif [9]. Teknik kualitatif merupakan peramalan berdasarkan pendapat suatu pihak, dan datanya tidak dapat dipresentasikan secara tegas menjadi suatu angka/nilai. Teknik peramalan tersebut misalnya adalah peramalan pendapat. Sebaliknya, teknik peramalan kuantitatif merupakan teknik peramalan yang berdasarkan pada data masa lalu (data historis) dan dapat dibuat dalam bentuk angka yang biasa disebut sebagai data *time series*.

2.2 Time Series

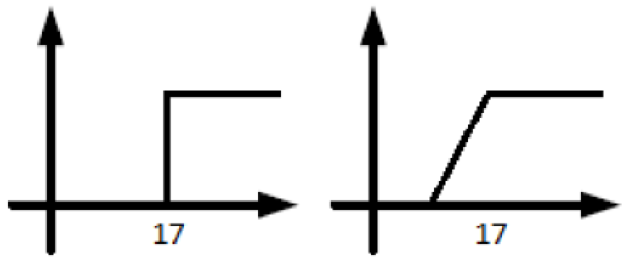
Time series merupakan jenis data yang dikumpulkan menurut urutan waktu dalam suatu rentang waktu tertentu. Waktu yang digunakan biasanya berupa hari, bulan, tahun, dan sebagainya. Analisis data berkala adalah salah satu prosedur statstika yang diterapkan untuk meramalkan struktur probabilitas keadaan yang akan datang dalam rangka pengambilan keputusan [10]. Dalam analisis deret waktu langkah awal yang biasanya dilakukan adalah menggambarkan nilai-nilai variabel terdahulu yang hendak diprediksikan pada sumbu vertikal dan waktu pada sumbu horizontal yang digunakan untuk menyelidiki secara visual gerakan deret waktu pada suatu jangka waktu tertentu. Selain hal itu, analisis deret waktu mencoba memprediksikan nilai-nilai di waktu yang akan datang dari deret waktu dengan mengkaji beberapa observasi data yang lalu [11].

2.3 Teori Himpunan *Fuzzy*

Teori himpunan *fuzzy* pertama kali diperkenalkan oleh Lotfi Asker Zadeh pada tahun 1965. Zadeh memperluas teori mengenai himpunan klasik menjadi himpunan kabur (*fuzzy* set) sehingga himpunan klasik (*crisp* set) merupakan kejadian khusus dari himpunan *fuzzy*. Himpunan *fuzzy* adalah himpunan yang anggotanya memiliki derajat keanggotaan tertentu yang nilainya berada pada selang tertutup [0,1]. Pada himpunan *fuzzy* nilai keanggotaan terletak pada rentang 0 sampai 1, yang berarti himpunan *fuzzy* dapat mewakili interpretasi tiap nilai berdasarkan pendapat atau keputusan probabilitasnya. Nilai 0 menunjukkan salah dan nilai 1 menunjukkan benar dan masih ada nilai-nilai yang terletak antara benar dan salah, dengan kata lain kebenaran suatu item tidak hanya benar atau salah [12].

Himpunan *crisp* adalah himpunan klasik yang telah dikenal secara umum. Himpunan *crisp* membedakan anggota hanya dengan nilai nol dan satu, anggota himpunan atau bukan [12]. Sebagai contoh himpunan yaitu, himpunan manusia. Himpunan wanita dan himpunan laki-laki dapat di representasikan dengan mudah dengan cara himpunan klasik. Akan tetapi, bagaimana mempresentasikan himpunan pada manusia muda atau tua. Muda atau tua itu cukup relative tidak langsung terpisah hanya karena berbeda umur satu hari. Dalam hal ini himpunan *fuzzy* dapat dikelompokkan dengan memberi nilai derajat tertentu. Berbeda dengan himpunan klasik, keanggotaan himpunan *fuzzy* dapat bernilai parsial.

Misalnya dalam kehidupan sehari-hari, usia dewasa didefenisikan dengan berusia 17 tahun keatas. Jika menggunakan himpunan *crisp,* seseorang yang berusia 17 tahun kurang 1 hari akan didefenisikan sebagai tidak dewasa. Namun, dalam himpunan *fuzzy,* orang tersebut dapat dinyatakan dengan hampir dewasa.



Gambar 2. 1 *Himpunan* crisp *(kiri) dan himpunan* fuzzy *(kanan)*

Fungsi keanggotaan (*membership function*) adalah suatu kurva yang menunjukkan pemetaan titik-titik input data ke dalam nilai keanggotaannya (derajat keanggotaan) yang memiliki interval antara 0 sampai 1 [13]. Fungsi keanggotaan didefenisikan sebagai berikut: jika X adalah himpunan semesta, maka fungsi keanggotaan μ_A (fungsi keanggotaan/fungsi karakteristik A pada X) yang didefenisikan oleh himpunan *fuzzy A* memiliki ketentuan sebagai berikut :

$$\mu_{\boldsymbol{A}}: X \rightarrow [0,1]$$

Dimana [0,1] adalah interval bilangan real dari nol sampai satu. Nilai x merupakan anggota penuh dari himpunan *fuzzy A*. Sementara nilai antara nol hingga satu menunjukkan bahwa x merupakan anggota himpunan *fuzzy A* secara parsial.

2.4 Fuzzy Time Series

Fuzzy Time Series (FTS) adalah metode prediksi yang menggunakan konsep *fuzzy* set sebagai dasar perhitungannya. Sistem prediksi dengan menggunakan metode ini dilakukan dengan menangkap pola dari data masa lalu kemudian digunakan untuk memproyeksikan data yang akan datang. Prosesnya juga tidak membutuhkan suatu sistem pembelajaran dari suatu sistem yang rumit, sebagaimana yang ada pada algoritma genetika dan jaringan syaraf sehingga mudah untuk digunakan dan dikembangkan.

Fuzzy Time Series (FTS) merupakan sebuah konsep yang diusulkan oleh Song dan Chissom dan digunakan untuk penyelesaian masalah prediksi dengan data historis berupa nilai-nilai linguistik [14]. Konsep dasar dari perhitungan model FTS adalah jika U merupakan himpunan semesta dengan $U = \{u_1, u_2, u_3, \dots, u_n\}$ maka himpunan *fuzzy* $A_i(i = 1,2,3,4, \dots, n)$ dapat di defenisikan sebagai berikut:

$$A_i = {f_{A_i}(u_1)}/{u_1} + {f_{A_i}(u_2)}/{u_2} + \cdots + {f_{A_i}(u_k)}/{u_k} \qquad (2.1)$$

Dimana f_{A_i} merupakan fungsi keanggotaan dari himpunan *fuzzy* A_i, u_k adalah elemen dari himpunan *fuzzy* A_i dan $f_{A_i}(u_k)$ adalah derajat keanggotaan dari u_k pada himpunan *fuzzy* A_i, k = 1,2,3, … , n.

Defenisi 2.4.1. Misalkan himpunan semesta $Y(t)(t = \cdots, 0,1,2, \dots, n, \dots)$, adalah subset dari R yang didefenisikan dengan himpunan *fuzzy* $A_{i(t)}$. Jika $F_{(t)}$ terdiri dari $A_{i(t)}$ ($i = 1,2, \dots, n$), maka $F_{(t)}$ didefenisikan sebagai FTS pada $Y(t)(t = \cdots, 0,1,2, \dots, n, \dots)$ [15].

Defenisi 2.4.2. Model peramalan FTS orde pertama musiman. Diberikan $F_{(t)}$ sebagai FTS. Asumsikan terdapat musiman di $\{F_{(t)}\}$, maka model peramalan *Fuzzy Time Series* adalah : $F_{(t-m)} \rightarrow F_{(t)}$, dimana m adalah periode [15].

Defenisi 2.4.3. Andaikan $F_{(t)} = A_i$ disebabkan oleh $F_{(t-m)} = A_j$, maka *Fuzzy Logical Relationship* (FLR) didefenisikan sebagai $A_i \rightarrow A_j$ [16].

Defenisi 2.4.4. Jika terdapat FLR yang diperoleh dari *state* A_2, maka transisi dibuat ke *state* yang lain $A_j, j = 1,2, \dots, n$ seperti $A_2 \rightarrow A_3, A_2 \rightarrow A_2, A_2 \rightarrow A_1$. Oleh karena itu FLR dikelompokkan menjadi *Fuzzy Logical Relationship Group* (FLRG) seperti berikut [16] :

$$A_2 \rightarrow A_1, A_2, A_3 \qquad (2.2)$$

Berikut adalah langkah-langkah dalam penyelesaian model *Fuzzy Time Series* [16] :

Langkah 1. Menentukan himpunan semesta U, dengan U adalah data historis. Himpunan semesta U dapat didefenisikan dengan $[D_{min} - D_1;\ D_{max} + D_2]$ dimana D_{min} merupakan nilai minimum pada data historis, D_{max} nilai maksimum pada data historis, dan D_1 dan D_2 adalah bilangan positif yang sesuai.

Langkah 2. Membagi himpunan semesta U menjadi beberapa interval (k) dengan rentang interval yang sama dengan menggunakan rumus *Sturges* berikut:

$$k = 1 + 3{,}322 \log N \tag{2.3}$$

dengan N adalah banyaknya data historis.

Perbedaan antara dua interval berturut-turut dapat didefenisikan dengan panjang interval (l) sebagai berikut:

$$l = \frac{(D_{max}+D_2)-(\ D_{min}-D_1)}{k} \tag{2.4}$$

Sehingga setiap interval (k) yang diperoleh adalah sebagai berikut :

$$u_1 = [D_{min} - D_1; D_{min} - D_1 + l]$$

$$u_2 = [D_{min} - D_1 + l; D_{min} - D_1 + 2l]$$

$$\vdots$$

$$u_n = [D_{min} - D_1 + (n-1)l; D_{min} - D_1 + nl] \tag{2.5}$$

Setelah himpunan semesta U telah dipartisi dengan diperolehnya banyak interval dan panjang tiap intervalnya, maka dapat dihitung nilai tengah dari masing-masing interval yang dapat didefenisikan dengan persamaan berikut :

$$m_n = \frac{(batas\ bawah + batas\ atas)}{2} \tag{2.6}$$

Langkah 3. Menentukan himpunan *fuzzy* untuk seluruh himpunan semesta U. Tidak ada batasan dalam menentukan banyaknya variabel linguistik yang dapat dijadikan himpunan *fuzzy*. Untuk setiap himpunan *fuzzy* A_i $(i = 1,2,3, \dots, n)$ didefenisikan dalam jumlah n interval yaiu $u_1 = [d_1; d_2], u_2 = [d_2; d_3], u_3 = [d_3; d_4], \dots, u_n = [d_n; d_{n+1}]$.

Adapun seluruh himpunan *fuzzy* dapat didefenisikan sebagai berikut [17] :

$$A_1 = \left\{ {}^{1}/_{u_1} + {}^{0,5}/_{u_2} + {}^{0}/_{u_3} + {}^{0}/_{u_4} + {}^{0}/_{u_5} + \cdots + {}^{0}/_{u_n} \right\}$$

$$A_2 = \left\{ {}^{0,5}/_{u_1} + {}^{1}/_{u_2} + {}^{0,5}/_{u_3} + {}^{0}/_{u_4} + {}^{0}/_{u_5} + \cdots + {}^{0}/_{u_n} \right\}$$

$$\vdots$$

$$A_n = \left\{ {}^{0}/_{u_1} + {}^{0}/_{u_2} + {}^{0}/_{u_3} + {}^{0}/_{u_4} + {}^{0}/_{u_5} + \cdots + {}^{0,5}/_{u_{n-1}} + {}^{1}/_{u_n} \right\} \quad (2.7)$$

Adapun aturan dalam menentukan derajat keanggotaan u_i adalah sebagai berikut :

$$A_i = \sum_{j=1}^{n} {}^{u_{ij}}/_{u_{ij}} \quad (2.8)$$

dengan u_{ij} adalah derajat keanggotaan u_{ij} milik A_i yang ditentukan sebagai berikut :

$$u_{ij} = \begin{cases} 1 \,;\, i = j \\ 0{,}5 \,;\, j = i - 1 \; atau \; i = j - 1 \\ 0 \,;\, lainnya \end{cases} \quad (2.9)$$

Berikut adalah beberapa aturan :

Aturan 1. Jika data historis (Y_t) adalah u_i, maka derajat keanggotaan u_i adalah 1, u_{i+1} adalah 0.5 dan lainnya adalah 0.

Aturan 2. Jika data historis (Y_t) adalah u_i, $1 < i < n$, maka derajat keanggotaan u_i adalah 1, u_{i+1} adalah 0.5 dan lainnya adalah 0.

Aturan 3. Jika data historis (Y_t) adalah u_n, maka derajat keanggotaan u_n adalah 1, u_{n-1} adalah 0.5 dan lainnya adalah 0.

Langkah 4. Melakukan fuzzifikasi, yaitu mengubah variabel-variabel numerik menjadi variabel linguistik pada data historis. Langkah ini bertujuan untuk menentukan himpunan *fuzzy* yang sesuai untuk setiap data.

Langkah 5. Menentukan FLR dan FLRG

Langkah 6. Menghitung nilai *output* yang akan diprediksikan. Jika $F(t-1) = A_j$, dan prediksi dari $F(t)$ dapat ditentukan dengan aturan-aturan dasar sebagai berikut :

Aturan 1. Jika FLRG dari A_j adalah himpunan kosong ($A_j \rightarrow \emptyset$), maka peramalan dari $F(t)$ adalah m_j, dimana m_j merupakan titik tengah dari interval u_j :

$$F(t) = m_j \tag{2.10}$$

Aturan 2. Jika FLRG dari A_j adalah himpunan satu ke satu ($A_j \rightarrow A_{k,} j, k = 1,2, \dots n$), maka peramalan dari $F(t)$ adalah m_k, yaitu titik tengah dari interval u_k :

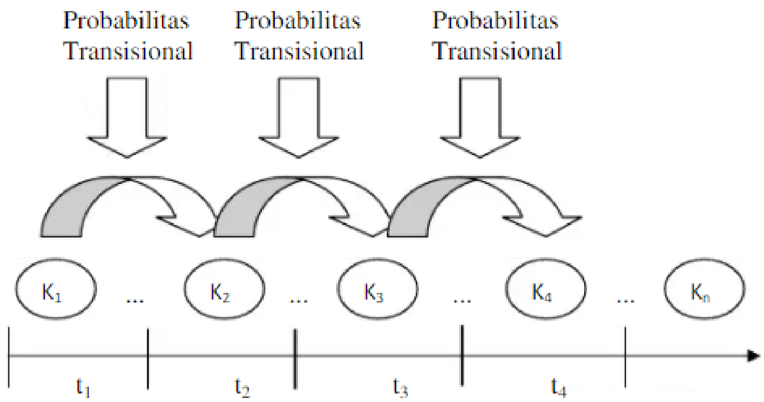
$$F(t) = m_k \tag{2.11}$$

Aturan 3. Jika FLRG dari $A_j$adalah satu ke banyak ($A_j \rightarrow A_{1,} A_3, A_5, j = 1,2, \dots, n$), maka peramalan $F(t)$ adalah sama untuk penghitungan rata-rata dari m_1, m_3, m_5, titik tengah interval u_1, u_3, u_5 :

$$F(t) = \frac{(m_1 + m_2 + \cdots + m_n)}{n} \tag{2.12}$$

2.5 Rantai Markov (Markov Chain)

Markov Chain adalah suatu metode yang mempelajari sifat-sifat suatu variabel pada masa sekarang yang didasarkan pada sifat-sifat masa lalu dalam usaha menaksir sifat-sifat variabel tersebut di masa yang akan datang. Teknik ini pertama kali dikembangkan oleh seorang ahli Rusia bernama A. A. Markov pada tahun 1906 [18]. A.A. Markov menejelaskan dalam temuannya bahwa: “Untuk setiap waktu t, ketika kejadian adalah K_t, dan seluruh kejadian sebelumnya adalah $K_{t(j)}, \dots, K_{t(j-n)}$ yang terjadi dari proses yang diketahui. Probabilitas seluruh kejadian yang akan datang $K_{t(j)}$ hanya bergantung pada kejadian $K_{t(j-1)}$ dan tidak bergantung pada kejadian-kejadian sebelumnya yaitu $K_{t(j-2)}, K_{t(j-3)}, \dots, K_{t(j-n)}$. "



Gambar 2. 2 Peristiwa dalam Rantai Markov

Jika $X_n = i$, maka proses ini terjadi di i pada saat n. Dengan menganggap bahwa kapan pun proses ini terjadi di *state* i, terdapat sebuah titik peluang P_{ij} yang akan berpindah ke *state* j [14]. Dengan demikian dapat dituliskan sebagai berikut :

$$P\{X_{n+1} = j | X_n = i, X_{n+1} - i_{n-1}, \dots, X_1 = i_1, X_0 = i_0\} = P_{ij} \quad (2.13)$$

untuk semua *state* $i_0, i_1, \dots, i_{n-1}, i, j, n \geq 0$. Proses yang seperti itu disebut rantai *Markov*.

Jika antara *state* A_i membuat transisi dengan *state* $A_j$dan melewati *state* lainnya $A_k, i, j = 1,2,3, \dots, n$ maka dapat diperoleh FLRG. Penentuan FLRG yang merupakan pengelompokan dari setiap perpindahan *state*, yaitu *state* saat ini (*current state*) dan *state* selanjutnya (*next state*). Pada setiap FLRG terdapat hubungan antara dua *state* yang disebut dengan *current state* dan *next state. Current state* merupakan nilai yang akan dihitung sebagai nilai prediksi. Sedangkan *next state* merupakan data yang digunakan sebagai syarat untuk memperoleh nilai pada *current state* [14].

Probabilitas transisional untuk *state* tersebut dapat ditulis sebagai berikut :

$$P_{ij} = \frac{M_{ij}}{M_i}, i, j = 1,2,3, \dots, n \qquad (2.14)$$

dengan,

P_{ij} adalah probabilitas transisional dari *state* A_i ke A_j dengan satu langkah.

M_{ij} adalah waktu transisional dari *state* A_i ke A_j dengan satu langkah.

M_i adalah jumlah data dari *state* A_i.

Sehingga matriks probabilitas transisional R dapat ditulis sebagai berikut :

$$R = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix} \qquad (2.15)$$

dari matriks R, terdapat beberapa defenisi yang dikembangkan, yaitu:

Defenisi 1. Jika $P_{ij} \geq 0$, maka *state* A_j dapat diakses dari *state* A_i.

Defenisi 2. Jika *state* A_i dan A_j saling dapat diakses, maka A_i berkomunikasi dengan A_j [16].

2.6 Model Fuzzy Time Series – Markov Chain

Langkah-langkah prediksi dari Langkah 1 sampai Langkah 5 pada model *Fuzzy Time Series – Markov Chain* (FTS-MC) adalah sama dengan model *Fuzzy Time Series* (FTS). Namun, yang membedakan antara model FTS dan FTS-*Markov Chain* yaitu pada langkah 6 sampai dengan langkah 8 sebagai berikut [16] :

Langkah 6. Menghitung nilai prediksi awal. Untuk data *time series*, digunakan *Fuzzy Logical Relationship Group* (FLRG) yang dapat digunakan untuk mendapatkan probabilitas *state* selanjutnya. Sehingga diperoleh matriks transisi *Markov*, *state* n

didefenisikan untuk setiap langkah waktu himpunan *fuzzy* n sehingga dimensi dari matriks transisi adalah $n \ x \ n$.

Matriks probabilitas yang telah diperoleh pada tahap sebelumnya dapat digunakan untuk menghitung nilai prediksi awal dengan aturan sebagai berikut :

Aturan 1. Jika FLRG A_i adalah kosong ($A_i \rightarrow$ Ø) maka hasil peramalan $F(t)$ adalah m_i, yaitu nilai tengah dari u_i dengan persamaan :

$$F(t) = m_i \tag{2.16}$$

Aturan 2. Jika FLRG A_i adalah himpunan satu ke satu ($A_i \rightarrow A_k$ dengan $P_{ij} = 0$ dan $P_{ik} = 1, j \neq k$), maka hasil prediksi $F(t)$ adalah m_k yaitu nilai tengah dari u_k dengan persamaan :

$$F(t) = m_k P_{ik} = m_k \tag{2.17}$$

Aturan 3. Jika FLRG A_i adalah himpunan satu ke banyak ($A_j \rightarrow A_{1,} A_3, A_5, j = 1,2, \dots, n$), jika kumpulan data $Y_{(t-1)}$ pada saat t-1 yang berada pada *state* A_j, maka hasil prediksi $F(t)$ adalah sebagai berikut :

$$F(t) = m_1 P_{j1} + m_2 P_{j2} + \cdots + m_{j-1} P_{j(j-1)} + Y_{(t-1)} P_j + m_{j+1} P_{j(j+1)} + \cdots + m_n P_{jn} \tag{2.18}$$

Dengan $m_1, m_2, \dots, m_{j-1}, m_{j+1}, \dots, m_n$ merupakan titik tengah dari $u_1, u_2, \dots, u_{j-1}, u_{j+1}, \dots, u_n$ dan m_j disubtitusikan ke $Y_{(t-1)}$ agar diperoleh informasi dari *state* A_j saat $t - 1$.

Langkah 7. Menyelesaikan kecenderungan nilai prediksi. Untuk percobaan *time series*, sampel berukuran besar selalu dibutuhkan. Oleh karena itu, ukuran sampel yang kecil ketika memodelkan model FTS-MC diperoleh matriks *Markov Chain* yang bias, dan beberapa penyesuaian untuk memprediksi nilai disarankan untuk meninjau kembali kesalahan prediksi. Aturan penyesuaian untuk nilai prediksi dijelaskan sebagai berikut :

Aturan 1. Jika *state* A_i berkomunikasi dengan A_j, dimulai dari *state* A_i pada saat $t-1$ sebagaimana $F_{(t-1)} = A_i$ dan terjadi perpindahan transisi naik ke *state* A_j pada saat t, $(i < j)$, maka nilai penyesuaian D_t ditentukan sebagai berikut :

$$D_{t1} = \left(\frac{l}{2}\right) \tag{2.19}$$

Aturan 2. Jika *state* A_i berkomunikasi dengan A_j, dimulai dari *state* A_i pada saat $t-1$ sebagaimana $F_{(t-1)} = A_i$ dan terjadi perpindahan transisi turun ke *state* A_j pada saat t, $(i > j)$, maka nilai penyesuaian D_t ditentukan sebagai berikut :

$$D_{t1} = -\left(\frac{l}{2}\right) \tag{2.20}$$

Aturan 3. Jika *state* A_i pada saat $t-1$ sebagaimana $F_{(t-1)} = A_i$ dan terjadi perpindahan transisi maju ke *state* A_{i+s} pada saat $t, 1 \leq s \leq n-i$, maka nilai penyesuaian D_t ditentukan sebagai berikut :

$$D_{t2} = \left(\frac{l}{2}\right)s \tag{2.21}$$

dengan s adalah banyak perpindahan transisi maju.

Aturan 4. Jika *state* A_i pada saat $t-1$ sebagaimana $F_{(t-1)} = A_i$ dan terjadi perpindahan transisi mundur ke *state* A_{i-v} pada saat $t, 1 \leq v \leq i$, maka nilai penyesuaian D_t ditentukan sebagai berikut :

$$D_{t2} = -\left(\frac{l}{2}\right)v \tag{2.22}$$

dengan v adalah banyak perpindahan transisi mundur.

Langkah 8. Menentukan hasil prediksi akhir. Bentuk umum untuk hasil prediksi akhir $F'(t)$ adalah sebagai berikut :

$$F'(t) = F(t) \pm D_{t1} \pm D_{t2} = F(t) \pm \frac{l}{2} \pm \left(\frac{l}{2}\right) v \qquad (2.23)$$

dengan l adalah rata-rata dari selisih interval yang berurutan dan v adalah perpindahan transisi.

2.7 Perhitungan *Error*

Perhitungan *error* merupakan suatu cara untuk mengetahui ketepatan model yang telah diperoleh. Perhitungan ini dapat diketahui seberapa akurat data hasil prediksi dari model yang digunakan dengan data aktualnya. Semakin kecil nilai *error* yang diperoleh maka semakin baik model prediksi yang digunakan. Salah satu metode perhitungan *error* ini adalah dengan menggunakan *Mean Absolute Percentage Error* (MAPE). MAPE merupakan rata-rata dari keseluruhan presentase kesalahan (selisih) antara data aktual dengan data hasil prediksi [16]. Berikut rumus metode MAPE :

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \frac{|Y(t) - F'(t)|}{Y(t)} x \; 100\% \qquad (2.24)$$

dengan $Y(t)$ merupakan data aktual, dan $F'(t)$ merupakan data hasil prediksi

Kriteria keakuratan dari metode perhitungan *error* MAPE ini dijelaskan pada tabel berikut :

Tabel 2. 1 Kriteria tingkat keakuratan MAPE

Kriteria Prediksi	**Batas Presentase MAPE**
Prediksi sangat baik	MAPE < 10%
Prediksi baik	MAPE 10% - 20%
Prediksi cukup	MAPE 20% - 50%
Prediksi tidak akurat	MAPE > 50%